## Title

Multiethnic Genome-Wide Association Study of Diabetic Retinopathy Using Liability Threshold Modeling of Duration of Diabetes and Glycemic Control.



## Journal

Diabetes, 68(2)

## ISSN

0012-1797

## Authors

Pollack, Samuela
Igo, Robert P
Jensen, Richard A
et al.

## Publication Date

2019-02-01

## DOI

10.2337/db18-0567

Peer reviewed

# Multiethnic Genome-Wide Association Study of Diabetic Retinopathy Using Liability Threshold Modeling of Duration of Diabetes and Glycemic Control

Samuela Pollack,[1] Robert P. Igo Jr.,[2] Richard A. Jensen,[3] Mark Christiansen,[3] Xiaohui Li,[4] Ching-Yu Cheng,[5,6] Maggie C.Y. Ng,[7,8] Albert V. Smith,[9] Elizabeth J. Rossin,[10] Ayellet V. Segrè,[10] Samaneh Davoudi,[10] Gavin S. Tan,[5,6] Yii-Der Ida Chen,[4] Jane Z. Kuo,[4,11] Latchezar M. Dimitrov,[7,8] Lynn K. Stanwyck,[10] Weihua Meng,[12] S. Mohsen Hosseini,[13] Minako Imamura,[14,15,16] Darryl Nousome,[17] Jihye Kim,[18] Yang Hai,[4] Yucheng Jia,[4] Jeeyun Ahn,[19] Aaron Leong,[20] Kaanan Shah,[21] Kyu Hyung Park,[22] Xiuqing Guo,[4] Eli Ipp,[23] Kent D. Taylor,[4] Sharon G. Adler,[24] John R. Sedor,[25,26,27] Barry I. Freedman,[28] Family Investigation of Nephropathy and Diabetes-Eye Research Group, DCCT/EDIC Research Group, I-Te Lee,[29,30,31] Wayne H.-H. Sheu,[29,30,31,32] Michiaki Kubo,[33] Atsushi Takahashi,[34,35] Samy Hadjadj,[36,37,38,39] Michel Marre,[40,41,42] David-Alexandre Tregouet,[43,44] Roberta Mckean-Cowdin,[17,45] Rohit Varma,[17,45] Mark I. McCarthy,[46,47,48] Leif Groop,[49] Emma Ahlqvist,[49] Valeriya Lyssenko,[49,50] Elisabet Agardh,[49] Andrew Morris,[51] Alex S.F. Doney,[52] Helen M. Colhoun,[53] Iiro Toppila,[54,55,56] Niina Sandholm,[54,55,56] Per-Henrik Groop,[54,55,56,57] Shiro Maeda,[14,15,16] Craig L. Hanis,[18] Alan Penman,[58] Ching J. Chen,[59] Heather Hancock,[59] Paul Mitchell,[60] Jamie E. Craig,[61] Emily Y. Chew,[62] Andrew D. Paterson,[63,64,65] Michael A. Grassi,[66,67] Colin Palmer,[68] Donald W. Bowden,[7,8] Brian L. Yaspan,[69] David Siscovick,[70] Mary Frances Cotch,[62] Jie Jin Wang,[5,60] Kathryn P. Burdon,[71] Tien Y. Wong,[5,72] Barbara E.K. Klein,[73] Ronald Klein,[73] Jerome I. Rotter,[4] Sudha K. Iyengar,[2] Alkes L. Price,[1] and Lucia Sobrin[10]

*Diabetes 2019;68:441–456 | https://doi.org/10.2337/db18-0567*

**To identify genetic variants associated with diabetic retinopathy (DR), we performed a large multiethnic genome-wide association study. Discovery included eight European cohorts (*n* = 3,246) and seven African American cohorts (*n* = 2,611). We meta-analyzed across cohorts using inverse-variance weighting, with and without liability threshold modeling of glycemic control and duration of diabetes. Variants with a *P* value $<1 \times 10^{-5}$ were investigated in replication cohorts that included 18,545 European, 16,453 Asian, and 2,710 Hispanic subjects. After correction for multiple testing, the C allele of rs142293996 in an intron of nuclear VCP-like (*NVL*) was associated with DR in European discovery cohorts ($P = 2.1 \times 10^{-9}$), but did not reach genome-wide significance after meta-analysis with replication cohorts. We applied the Disease Association Protein-Protein Link Evaluator (DAPPLE) to our discovery results to test for evidence of risk being spread across underlying molecular pathways. One protein–protein interaction network built from genes in regions associated with proliferative DR was found to have significant connectivity (*P* = 0.0009) and corroborated with gene set enrichment analyses. These findings suggest that genetic variation in *NVL*, as well as variation within a protein–protein interaction network that includes genes implicated in inflammation, may influence risk for DR.**

Diabetic retinopathy (DR) is a leading cause of blindness (1). Established risk factors include longer duration of

[1]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA
[2]Department of Population and Quantitative Health Sciences, Case Western University, Cleveland, OH
[3]Cardiovascular Health Research Unit, Department of Medicine, Epidemiology and Health Services, University of Washington, Seattle, WA
[4]Institute for Translational Genomics and Population Sciences, LA BioMed and Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, CA
[5]Duke-NUS Medical School, Singapore
[6]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore
[7]Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC
[8]Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC
[9]Department of Medicine, University of Iceland, Reykjavík, Iceland

diabetes (DoD) and poor glycemic control (2). Genetic factors are also implicated, with heritability of 52% for proliferative DR (PDR) (3,4). Several candidate gene and genome-wide association studies (GWAS) have been conducted (5–11). Although several polymorphisms have been suggested to be associated with DR, few have been convincingly replicated (10,12–15).

There are several reasons why studies have not yielded consistent findings. The genetic effects are likely modest, and identification requires large sample sizes. Previous studies have not consistently accounted for the strongest two covariates, DoD and glycemic control. Liability threshold (LT) modeling is one way to incorporate these covariates while also increasing statistical power (16). Finally,

[10]Massachusetts Eye and Ear Department of Ophthalmology, Harvard Medical School, Boston, MA
[11]Medical Affairs, Ophthalmology, Sun Pharmaceutical Industries, Inc., Princeton, NJ
[12]Division of Population Health Sciences, Ninewells Hospital and Medical School, University of Dundee School of Medicine, Scotland, U.K.
[13]Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada
[14]Laboratory for Endocrinology, Metabolism and Kidney Diseases, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan
[15]Department of Advanced Genomic and Laboratory Medicine, Graduate School of Medicine, University of the Ryukyus, Nishihara, Japan
[16]Division of Clinical Laboratory and Blood Transfusion, University of the Ryukyus Hospital, Nishihara, Japan
[17]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA
[18]Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX
[19]Department of Ophthalmology, SMG-SNU Boramae Medical Center, Seoul National University College of Medicine, Seoul, Korea
[20]Endocrine Unit and Diabetes Unit, Division of General Internal Medicine, Massachusetts General Hospital, Boston, MA
[21]Section of Genetic Medicine, University of Chicago, Chicago, IL
[22]Department of Ophthalmology, Seoul National University College of Medicine, Seoul National University Bundang Hospital, Seongnam, Korea
[23]Section of Diabetes and Metabolism, Harbor-UCLA Medical Center, University of California, Los Angeles, Los Angeles, CA
[24]Department of Nephrology and Hypertension, Los Angeles Biomedical Research Institute at Harbor-University of California, Torrance, CA
[25]Department of Medicine, Case Western Reserve University, Cleveland, OH
[26]Department of Physiology and Biophysics, Case Western Reserve University, Cleveland, OH
[27]Division of Nephrology, MetroHealth System, Cleveland, OH
[28]Section on Nephrology, Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC
[29]Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan
[30]School of Medicine, Chung Shan Medical University, Taichung, Taiwan
[31]School of Medicine, National Yang-Ming University, Taipei, Taiwan
[32]School of Medicine, National Defense Medical Center, Taipei, Taiwan
[33]RIKEN Center for Integrative Medical Sciences, Yokohama, Japan
[34]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Kanagawa, Japan
[35]Department of Genomic Medicine, Research Institute, National Cerebral and Cardiovascular Center, Osaka, Japan
[36]CHU de Poitiers, Centre d'Investigation Clinique, Poitiers, France
[37]Université de Poitiers, UFR Médecine Pharmacie, Centre d'Investigation Clinique 1402, Poitiers, France
[38]INSERM, Centre d'Investigation Clinique 1402, Poitiers, France
[39]L'Institut du Thorax, INSERM, CNRS, CHU Nantes, Nantes, France
[40]Université Paris Diderot, Sorbonne Paris Cité, Paris, France
[41]Department of Diabetology, Endocrinology and Nutrition, Assistance Publique–Hôpitaux de Paris, Bichat Hospital, DHU FIRE, Paris, France
[42]INSERM U1138, Centre de Recherche des Cordeliers, Paris, France
[43]Team Genomics & Pathophysiology of Cardiovascular Diseases, UPMC, Sorbonne Universités, INSERM, UMR_S 1166, Paris, France
[44]Institute of Cardiometabolism and Nutrition, Paris, France
[45]Department of Ophthalmology, USC Roski Eye Institute, Keck School of Medicine of the University of Southern California, Los Angeles, CA
[46]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, U.K.
[47]Wellcome Centre for Human Genetics, University of Oxford, Oxford, U.K.
[48]NIHR Oxford Biomedical Research Centre, Churchill Hospital, Oxford, U.K.
[49]Department of Clinical Sciences, Faculty of Medicine, Lund University, Malmö, Sweden
[50]Department of Clinical Science, KG Jebsen Center for Diabetes Research, University of Bergen, Bergen, Norway
[51]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, U.K.
[52]Molecular and Clinical Medicine, Ninewells Hospital and Medical School, University of Dundee, Dundee, U.K.
[53]Institute of Genetics and Molecular Medicine, Western General Hospital, University of Edinburgh, Edinburgh, U.K.
[54]Folkhälsan Institute of Genetics, Folkhälsan Research Center, Helsinki, Finland
[55]Abdominal Center, Nephrology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[56]Research Programs Unit, Diabetes and Obesity, University of Helsinki, Helsinki, Finland
[57]Department of Diabetes, Central Clinical School, Monash University, Melbourne, Victoria, Australia
[58]Department of Preventive Medicine, John D. Bower School of Population Health, University of Mississippi Medical Center, Jackson, MS
[59]Department of Ophthalmology, University of Mississippi Medical Center, Jackson, MS
[60]Centre for Vision Research, Westmead Institute for Medical Research, The University of Sydney, Sydney, New South Wales, Australia
[61]Department of Ophthalmology, Flinders University, Bedford Park, South Australia, Australia
[62]Division of Epidemiology and Clinical Applications, National Eye Institute, National Institutes of Health, Bethesda, MD
[63]Institute of Medical Sciences, University of Toronto, Toronto, Ontario, Canada
[64]Program in Genetics & Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada
[65]Epidemiology and Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
[66]Grassi Retina, Naperville, IL
[67]Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, IL
[68]Pat MacPherson Centre for Pharmacogenetics and Pharmacogenomics, Ninewells Hospital and Medical School, University of Dundee, Dundee, U.K.
[69]Genentech, Inc., South San Francisco, CA
[70]Institute for Urban Health, New York Academy of Medicine, New York, NY
[71]Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia

previous genetic studies have largely examined individual variants. Techniques that examine top GWAS findings collectively for variants that cluster in biological networks based on known protein–protein interactions have the potential to identify variants where there is insufficient power to detect their individual effects.

The purpose of this study was to identify genetic variants associated with DR by *1*) assembling a large sample size through inclusion of multiple ethnicities, *2*) incorporating DoD and glycemic control via LT modeling, and *3*) collectively examining variants that cluster in biological networks.

## RESEARCH DESIGN AND METHODS

All studies conformed to the Declaration of Helsinki tenets and were Health Insurance Portability and Accountability Act compliant. Written informed consent was obtained from all participants. Institutional Review Board/Ethics Committee approval was obtained by each individual study.

### Discovery Sample Description

The discovery sample, encompassing 7 African American and 8 European cohorts, arose from a consortium of 11 DR studies for a total of 3,246 Europeans and 2,611 African Americans (6–8,12,13,17,18). Inclusion criteria for the discovery stage were *1*) type 2 diabetes, and *2*) European or African American ethnicity. Type 2 diabetes was defined as a fasting plasma glucose (FPG) ≥126 mg/dL (7.0 mmol/L) or a hemoglobin $A_{1c}$ ($HbA_{1c}$) ≥6.5% (48 mmol/mol) (19) with onset of the diabetes after 30 years of age. Table 1 summarizes the DR phenotyping protocols and covariates by discovery cohort. Phenotyping protocols have been previously described (4,20–29), and additional details are in the Supplementary Data.

### DR Case-Control Definitions

The analysis plan prespecified four DR case-control definitions with varying Early Treatment Diabetic Retinopathy Study (ETDRS) score thresholds for case and control subjects (Table 2) (30). The primary case-control definition compared any DR to no DR (ETDRS ≥14 vs. ETDRS <14, henceforth referred to as the any DR analysis). There were three secondary case-control definitions. The first compared patients with PDR to those without PDR (ETDRS ≥60 vs. ETDRS <60, henceforth the PDR analysis). The second compared those with nonproliferative DR (NPDR) or worse to those without DR (ETDRS ≥30 vs. ETDRS <14, henceforth the NPDR analysis). The third compared those with PDR to those without DR (ETDRS ≥60 vs. ETDRS <14, henceforth the extremes of DR analysis). The rationale for the four definitions is in the Supplementary Data. Table 1 shows the available samples by cohort and ETDRS score thresholds. Supplementary Table 1 summarizes the mean values for glycemic control and DoD.

### Statistical Analyses

The genotyping platforms and numbers of single nucleotide polymorphisms (SNPs) genotyped are summarized in Supplementary Table 2. Details about quality control, imputation, and data filtering are in the Supplementary Data. Supplementary Fig. 1 provides a flowchart of the discovery and replication analyses. For the four main case-control definition analyses, we performed each of the analyses *1*) without incorporating DoD and glycemic control using EIGENSOFT (16,31) and *2*) with LT modeling of DoD and glycemic control using LTSCORE (16). LT modeling details are in the Supplementary Data. Both the EIGENSOFT and LTSCORE tests were implemented in LTSOFT version 2.0 (see Web Resources in the Supplementary Data). For the discovery analyses, we ran principal components (PC) analysis with EIGENSTRAT using only typed SNPs and five PCs, separately by ethnicity and case-control definition (32). We computed association analyses for each of the seven African American and eight European cohorts separately and then meta-analyzed by ethnicity. Meta-analysis was performed using inverse-variance weighting, accounting for both effective sample size (defined as $4/[1/N_{case} + 1/N_{control}]$) and allele frequency (33). We also performed multiethnic (Europeans and African Americans together) meta-analyses for the any DR and PDR analyses using inverse-variance weighting and a sensitivity analysis of the any DR meta-analyses in African Americans and Europeans (see Supplementary Data). Because we included rare variants in this GWAS, we also tested the robustness of the top associations ($P < 5 \times 10^{-8}$) by performing two additional tests: *1*) a Fisher exact test on case or control subjects aggregated across all cohorts tested per variant and on each cohort separately,

[72]Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[73]Department of Ophthalmology and Visual Sciences, University of Wisconsin-Madison, Madison, WI

Corresponding author: Lucia Sobrin, lucia_sobrin@meei.harvard.edu

Received 23 May 2018 and accepted 12 November 2018

This article contains Supplementary Data online at https://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db18-0567/-/DC1.

A.L.P. and L.S. contributed equally to this work.

S.M.H. is currently affiliated with the Department of Pathology and Laboratory Medicine, Hospital of the University of Pennsylvania, Philadelphia, PA.

C.J.C. is currently affiliated with the Retina Center, North Mississippi Medical Center, Tupelo, MS.

**Table 1—Studies included in the discovery sample**

| Study | Population | Diabetes type | Number of eyes/number of fields/size of fields photographed | Diabetes duration | Glycemic control measure | Case subjects (ETDRS ≥14) | Control subjects (ETDRS <14) | Case subjects (ETDRS ≥60) | Control subjects (ETDRS <60) | Case subjects (ETDRS ≥30) |
|---|---|---|---|---|---|---|---|---|---|---|
| AAPDR | AA | 2 | 2/7/30° | Y | $HbA_{1c}$ | 274 | 56 | 255 | 75 | 261 |
| AGES* | EUR | 2 | 2/2/45° | Y | $HbA_{1c}$ | 85 | 222 | 3 | 304 | 8 |
| ARIC | AA | 2 | 1/1/45° | Y | $HbA_{1c}$ | 96 | 265 | 3 | 358 | 73 |
| ARIC | EUR | 2 | 1/1/45° | Y | $HbA_{1c}$ | 126 | 632 | 6 | 752 | 80 |
| AUST | EUR | 2 | NA‡ | Y | $HbA_{1c}$ | 522 | 435 | 187 | 770 | 346 |
| BMES | EUR | 2 | 2/5/30° | Y | FPG | 124 | 208 | 1 | 331 | 37 |
| CHS | AA | 2 | 1/1/45° | Y | FPG | 19 | 35 | 4 | 50 | 14 |
| CHS | EUR | 2 | 1/1/45° | Y | FPG | 26 | 119 | 4 | 141 | 16 |
| FIND-Eye* | AA | 2 | 2/2/45°† | Y | $HbA_{1c}$ | 330 | 167 | 264 | 233 | 303 |
| FIND-Eye | EUR | 2 | 2/2/45°† | Y | $HbA_{1c}$ | 158 | 154 | 115 | 197 | 145 |
| JHS | AA | 2 | 2/7/30° | Y | $HbA_{1c}$ | 91 | 160 | 12 | 239 | 57 |
| MESA | AA | 2 | 2/2/45° | Y | $HbA_{1c}$ | 101 | 258 | 11 | 348 | 60 |
| MESA | EUR | 2 | 2/2/45° | Y | $HbA_{1c}$ | 38 | 200 | 2 | 236 | 12 |
| RISE/RIDE | EUR | 2 | 2/7/30° | Y | $HbA_{1c}$ | — | — | 80 | 117 | — |
| WFU | AA | 2 | NA‡ | Y | $HbA_{1c}$ | — | — | 548 | 211 | — |
| Total | AA | 2 | — | Y | Varies | 911 | 941 | 1,097 | 1,514 | 768 |
| Total | EUR | 2 | — | Y | Varies | 1,079 | 1,970 | 398 | 2,848 | 644 |

AA, African American; AAPDR, African American Proliferative Diabetic Retinopathy Study; AGES, Age, Gene/Environment, Susceptibility - Reykjavik Study; ARIC, Atherosclerosis Risk in Communities Study; AUST, Australian Genetics of Diabetic Retinopathy Study; BMES, Blue Mountains Eye Study; CHS, Cardiovascular Health Study; EUR, European; FIND-Eye, Family Study of Nephropathy and Diabetes-Eye; JHS, Jackson Heart Study; MESA, Multiethnic Study of Atherosclerosis; NA, not available; RIDE/RISE, Ranibizumab Injection in Subjects with Clinically Significant Macular Edema with Center Involvement Secondary to Diabetes; WFU, Wake Forest School of Medicine Study; Y, information on diabetes duration is available. *Cohorts without access to raw genotype information. †Not all FIND-Eye subjects had photographs, but all participants had harmonization of exam and clinical data to an ETDRS score. ‡AUST used examination by an ophthalmologist to ascertain DR. The WFU study used a questionnaire to ascertain DR.

**Table 2—Four case-control subject definitions and the number of samples available for discovery for each definition**

| | Control subjects | | | Case subjects | | |
|---|---|---|---|---|---|---|
| Analysis | Score | *n* AA | *n* EUR | Score | *n* AA | *n* EUR |
| Any DR (primary analysis) | <14 | 941 | 1,970 | ≥14 | 911 | 1,079 |
| PDR | <60 | 1,514 | 2,848 | ≥60 | 1,097 | 398 |
| NPDR | <14 | 941 | 1,970 | ≥30 | 768 | 644 |
| Extremes of DR | <14 | 941 | 1,970 | ≥60 | 1,097 | 398 |

AA, African American; EUR, European; Score, ETDRS score range.

and *2*) an inverse variance-weighted meta-analysis across cohorts using the ln of the odds ratio (OR) as the effect size (34) without adjusting for covariates.

### *P* Value Thresholds for Genome-Wide Significance

The *P* value thresholds for genome-wide significance were based on empirically determined thresholds for different ancestral populations that account for the GWAS multiple testing burden, as well as population-specific linkage disequilibrium (LD) patterns (35):

1. $P < 3.24 \times 10^{-8}$ for SNPs ascertained in African ancestry populations
2. $P < 5.0 \times 10^{-8}$ for SNPs ascertained in European ancestry populations
3. $P < 3.24 \times 10^{-8}$ for SNPs ascertained in multiethnic meta-analyses

We further corrected these thresholds for additional multiple testing from examination of four case-control definitions, each with and without covariate incorporation, for eight tests total. This yielded the following *P* value thresholds for our study:

4. $P < 3.75 \times 10^{-9}$ for SNPs ascertained in African ancestry populations
5. $P < 6.25 \times 10^{-9}$ for SNPs ascertained in European ancestry populations
6. $P < 3.75 \times 10^{-9}$ for SNPs ascertained in multiethnic meta-analyses

We note that correction for eight tests is conservative because the case-control definitions are not completely independent. We did not apply further multiple testing correction for the different ancestries analyzed.

### Replication Meta-Analysis

Eight European, eight Asian, and four Hispanic replication cohorts provided summary statistics on SNPs with $P < 1 \times 10^{-5}$ in the discovery analyses (Table 3). Their phenotyping/genotyping protocols have been previously described, and details are in the Supplementary Data (6–8,12,13,17,18). The rationale for including additional ethnicities in the replication phase is that high transethnic genetic correlations have been documented for type 2 diabetes and other traits/diseases and support the use of multiethnic studies to increase sample size (36). Supplementary Table 3 summarizes the replication cohorts' mean values for $HbA_{1c}$, FPG, and DoD. Replication was in silico with existing genotyping. LT modeling was not applied to the replication cohort analyses. The replication cohorts used standard covariate adjustment in their regression models. Replication meta-analysis was also performed using inverse-variance weighting, first individually by each ethnicity (Europeans, Hispanics, and Asians) followed by all cohorts combined. Replicated genome-wide significance had to meet the aforementioned thresholds after meta-analysis of the discovery and replication results.

### Protein–Protein Interaction Analysis of Top GWAS Loci

To identify significantly enriched protein networks among the loci with the highest statistical evidence for association with DR, we applied the Disease Association Protein-Protein Link Evaluator (DAPPLE) to our discovery GWAS (37). It has been shown that top associated loci, despite not being genome-wide significant, tend to cluster in biological networks (37,38). For this reason, we examined the top 1,000 loci from the discovery GWAS in the two monoethnic analyses (European and African American) and for each of the four case-control definition analyses that incorporated DoD and glycemic control (eight network analyses in total). Our threshold for significance was therefore $P < 0.00625$ (0.05 corrected for eight tests). We used the publically available version of DAPPLE, and the protocol is outlined in the Supplementary Data. This methodology has been used successfully with previous GWAS to identify protein networks with biological relevance (37–39).

### Gene Set Enrichment Analysis of DAPPLE Significant Genes

To further support the protein–protein interaction results from the DAPPLE analysis, we applied gene set enrichment analysis (GSEA) using Meta-Analysis Gene-Set Enrichment of variaNT Associations (MAGENTA) (40) to the set of genes significantly enriched for protein–protein interactions in the DAPPLE analysis (details in Supplementary Data).

### Type 2 Diabetes and Associated Glycemic Traits Loci

To understand to what extent genetic determination of DR might reflect enrichment for type 2 diabetes or glycemic

**Table 3—Studies included in the replication meta-analyses**

| Cohort by ancestry | Ethnicity/nationality | DM type | Any DR analysis | | PDR analysis | | NPDR analysis | | Extremes of DR analysis | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Case subjects | Control subjects | Case subjects | Control subjects | Case subjects | Control subjects | Case subjects | Control subjects |
| Asian | | | | | | | | | | |
| KSDR | Korean | 2 | 1,516 | 571 | 918 | 1,167 | 1,300 | 571 | 918 | 571 |
| MESA | Chinese | 2 | 28 | 83 | — | — | 17 | 83 | — | — |
| RIKEN | Japanese | 2 | 5,532 | 5,565 | — | — | 2,371 | 5,565 | — | — |
| SCES I | Chinese | 2 | 75 | 228 | — | — | — | — | — | — |
| SCES II | Chinese | 2 | 27 | 78 | — | — | — | — | — | — |
| SiMES | Malay | 2 | 214 | 557 | — | — | — | — | — | — |
| SINDI | Indian | 2 | 315 | 669 | — | — | — | — | — | — |
| TUDR | Chinese | 2 | — | — | — | — | — | — | 436 | 559 |
| European | | | | | | | | | | |
| DCCT/EDIC primary cohort | North American | 1 | — | — | 53 | 598 | — | — | — | — |
| DCCT/EDIC secondary cohort, conventional treatment | North American | 1 | — | — | 114 | 209 | — | — | — | — |
| DCCT/EDIC secondary cohort, intensive treatment | North American | 1 | — | — | 42 | 288 | — | — | — | — |
| GENESIS/GENEDIAB | French | 1 | 277 | 999 | 808 | 468 | 277 | 607 | 277 | 468 |
| GoDARTS | Scottish | | 2,506 | 2,412 | 574 | 4,345 | 1,381 | 2,412 | 574 | 2,412 |
| GoKinD | North American | 1 | — | — | 138 | 581 | — | — | — | — |
| SUMMIT | European | 1 and 2 | 5,422 | 4,302 | — | — | — | — | — | — |
| WESDR | North American | 1 | — | — | 309 | 294 | — | — | — | — |
| Hispanic | | | | | | | | | | |
| GOLDR | Hispanic | 2 | 298 | 301 | 76 | 523 | 215 | 301 | 76 | 301 |
| LALES | Hispanic | 2 | 552 | 500 | 53 | 999 | 341 | 500 | 53 | 500 |
| MESA | Hispanic | 2 | 92 | 192 | — | — | 52 | 192 | — | — |
| SCHS | Mexican American | 2 | 528 | 247 | 103 | 672 | 406 | 247 | 103 | 247 |
| Total | | | 17,382 | 16,704 | 3,188 | 10,144 | 6,360 | 10,478 | 2,437 | 5,058 |

The SUMMIT (SUrrogate markers for Micro- and Macrovascular hard endpoints for Innovative diabetes Tools) cohort is a meta-analysis of three European studies: the Finnish Diabetic Nephropathy (FinnDiane) Study, Scania Diabetes Registry, and the EURODIAB study. DCCT/EDIC, Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications; DM, diabetes mellitus; GENESIS/GENEDIAB, Genetics Nephropathy and Sib Pair Study/Génétique de la Nephropathie Diabétique; GoDARTS, Genetics of Diabetes and Audit Research Tayside Study; GoKinD, Genetics of Kidneys in Diabetes; GOLDR, Genetics of Latino Diabetic Retinopathy; KSDR, Korean Study of Diabetic Retinopathy; LALES, Los Angeles Latino Eye Study; MESA, Multiethnic Study of Atherosclerosis; RIKEN, Rikagaku Kenkyusho - Institute of Physical and Chemical Research; SCES, Singapore Chinese Eye Study; SCHS, Starr County Health Studies; SiMES, Singapore Malay Eye Study; SINDI, Singapore Indian Eye Study; TUDR, Taiwan–US Diabetic Retinopathy Study; WESDR, Wisconsin Epidemiologic Study of Diabetic Retinopathy.

control genes, we computed a correlation between case status in the any DR analysis and the sum of the β*risk allele (for quantitative glycemic traits) or logOR*risk allele (for type 2 diabetes) of the trait-associated SNPs for each cohort and each trait (see Supplementary Data for details).

## RESULTS

### Discovery Meta-analysis

Supplementary Fig. 2 shows the PC analysis. We observed little inflation in the association statistic distribution (Supplementary Fig. 3), indicating no significant population stratification as a confounder. Supplementary Fig. 4 shows the Manhattan plots for the any DR analyses. Supplementary Tables 4–25 show the top 10 SNPs for independent loci with the lowest *P* values for each discovery analysis, including the sensitivity analyses (full results are available on the Type 2 Diabetes Knowledge Portal [http://www.type2diabetesgenetics.org/], both on the downloads page and fully integrated into the portal modules).

Table 4 shows SNPs that met the traditional nominal threshold for genome-wide significance of $P < 5 \times 10^{-8}$ from the discovery analyses. All of the SNPs in Table 4 were either from the PDR or extremes of DR analyses; Fig. 1 shows the QQ and Manhattan plots for the PDR and extremes of DR analyses. The results for the associations in Table 4 are shown for each cohort separately in Supplementary Table 26. Results for these SNPs after meta-analysis with replication samples both combined and separated by ethnicity are shown in Table 5 and Supplementary Table 27, respectively.

### Genome-Wide Significant Finding From the Discovery Analyses in NVL Gene

Using the corrected significance thresholds, only one SNP in the discovery meta-analyses met genome-wide significance: rs142293996 for the extremes of DR analysis incorporating DoD and glycemic control in Europeans ($P = 2.1 \times 10^{-9}$). The association was not significant without adjusting for covariates based on a Fisher exact test (Supplementary Table 28). This is an intronic variant in the nuclear VCP-like (*NVL*) gene, which encodes a member of the ATPases associated with diverse cellular activities (AAA) superfamily (41). The *NVL* gene is widely expressed in vivo with highest expression in retina (https://www.proteinatlas.org/ENSG00000143748-NVL/tissue#top).

We tested whether this association was a significant *cis*-expression quantitative trait locus (eQTL) in the Genotype-Tissue Expression (GTEx) Project release v7 (see Supplementary Data for eQTL analysis details). This variant, rs142293996, lies in the 22nd intron of *NVL* and is in LD ($r^2 = 0.62$) with variant rs41271487 in the 24th intron of *NVL*. rs41271487 is a significant eQTL ($P = 6.4 \times 10^{-6}$; effect size 1.27) in the GTEx spinal cord cervical c-1 tissue, targeting calpain 2 (*CAPN2*), a calcium-activated neutral protease (Supplementary Fig. 5). Common variants in the intron or regulatory region of *CAPN2*, 527–576 kb upstream of the DR association, are associated with

**Table 4—Variants with $P < 5 \times 10^{-8}$ (traditional, nominal threshold for genome-wide significance) in the discovery analyses**

| Case-control definition | Population/LT modeling | RSID | CHR | Position | Nearest gene | REF | Case subjects | | Control subjects | | NEFF | *P* | OR | 95% CI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | *N* | RAF | *N* | RAF | | | | |
| PDR | AA/no | rs115523882 | 3 | 167876205 | *GOLIM4* | A | 1,105 | 0.9823 | 1,119 | 0.9611 | 1,452 | $9.42 \times 10^{-9}$ | 3.10 | 2.12, 4.53 |
| PDR | AA/yes | rs115523882 | 3 | 167876205 | *GOLIM4* | A | 1,105 | 0.9823 | 1,119 | 0.9611 | 1,452 | $5.37 \times 10^{-9}$ | 3.10 | 2.14, 4.50 |
| PDR | EUR/no | rs139205645 | 2 | 201949806 | *NDUFB3* | T | 309 | 0.9725 | 975 | 0.9959 | 907 | $3.93 \times 10^{-8}$ | 0.13 | 0.06, 0.27 |
| PDR | EUR/yes | rs17791488 | 17 | 26232732 | *NOS2/LYRM9* | T | 309 | 0.9871 | 975 | 0.9661 | 907 | $7.26 \times 10^{-9}$ | 3.70 | 2.40, 5.71 |
| Extremes of DR | AA/no | rs184340784 | 1 | 4589883 | *AJAP1* | C | 520 | 0.999 | 230 | 0.9784 | 603 | $3.52 \times 10^{-8}$ | NA | NA |
| Extremes of DR | EUR/yes | rs142293996 | 1 | 224448059 | *NVL* | C | 187 | 0.9947 | 435 | 0.9874 | 523 | $2.10 \times 10^{-9}$ | 2.38 | 1.80, 3.14 |
| Extremes of DR | EUR/yes | rs17706958 | 3 | 73837141 | *PDZRN3* | T | 308 | 0.8139 | 594 | 0.7332 | 797 | $3.04 \times 10^{-8}$ | 1.58 | 1.35, 1.85 |
| Extremes of DR | EUR/yes | rs80117617 | 2 | 40855125 | *SLC8A1* | T | 308 | 0.9838 | 594 | 0.9445 | 797 | $4.04 \times 10^{-8}$ | 3.78 | 2.37, 6.02 |

AA, African American; CHR, chromosome; EUR, European; LT, liability threshold; NA, not available; NEFF, effective sample size; RAF, reference allele frequency; REF, reference allele; RSID, rs identifier.

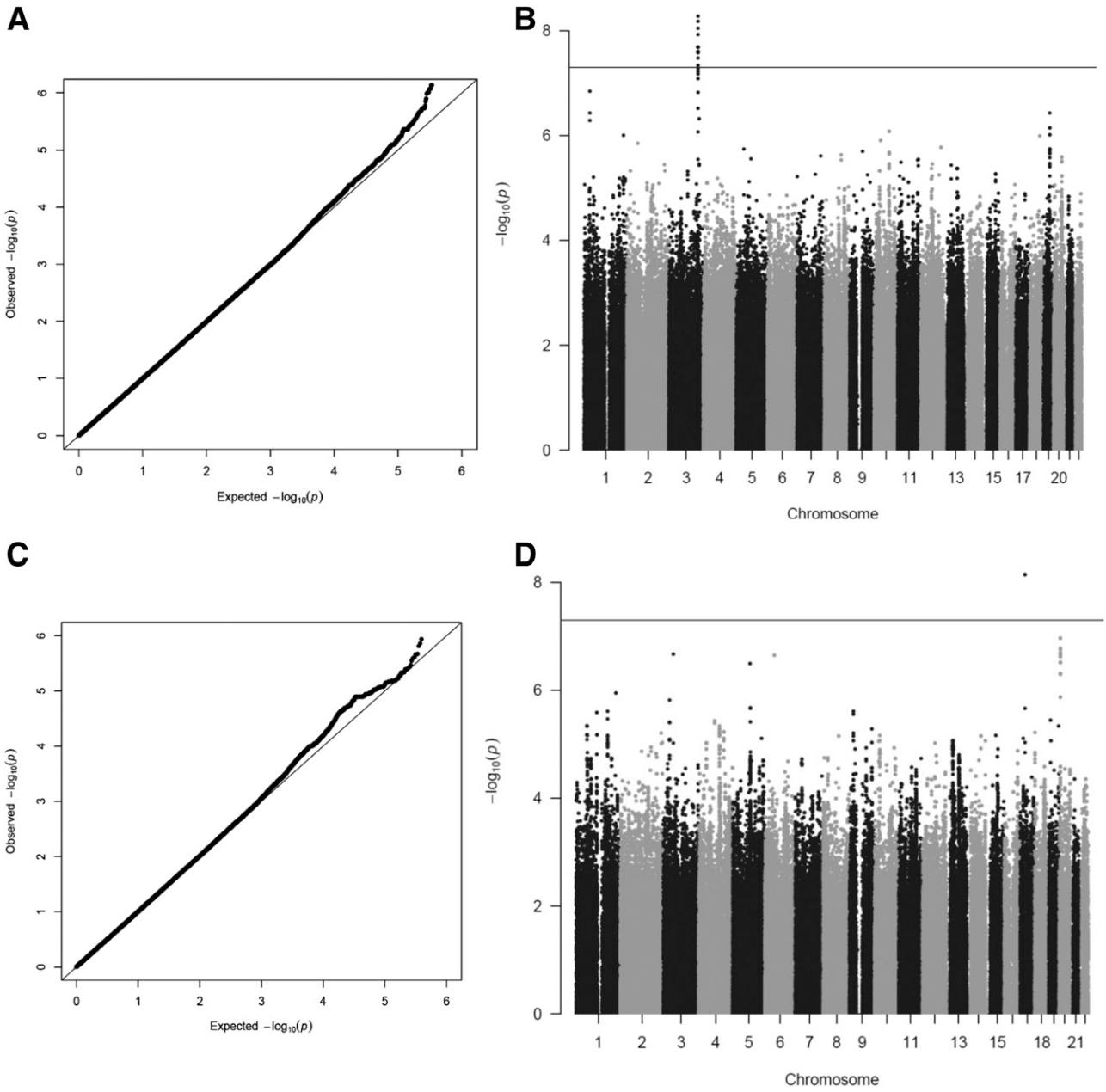


**Figure 1**—Quantile-quantile and Manhattan plots for the PDR and extremes of DR discovery meta-analyses for PDR analysis in African American participants with LT modeling of DoD and glycemic control (*A* and *B*), PDR analysis in European participants with LT modeling of DoD and glycemic control (*C* and *D*), extremes of DR analysis in African American participants with LT modeling of DoD and glycemic control (*E* and *F*), and extremes of DR analysis in European participants with LT modeling of DoD and glycemic control (*G* and *H*). The horizontal line in each of the Manhattan plots indicates the nominal threshold for genome-wide significance ($P = 5 \times 10^{-8}$).

variation in serum α-carotene levels (42), a vitamin A precursor required for sight, supporting a functional role for this gene. Based on the eQTL analysis, increased expression of *CAPN2* is associated with decreased risk of DR (Supplementary Fig. 6). *CAPN2* is expressed in the retina (https://www.proteinatlas.org/ENSG00000162909-CAPN2/tissue).

When examined in the replication analyses (which included a more diverse population), the direction of effect in the replication cohorts for rs142293996 was the same, but the meta-analysis *P* value was not genome-wide significant ($P = 4.10 \times 10^{-6}$).

### Top Finding From the African American Discovery Analyses

In African Americans, the SNP with the lowest *P* value was rs115523882 from the PDR analysis ($P = 5.37 \times 10^{-9}$). This was short of the $3.75 \times 10^{-9}$ threshold for

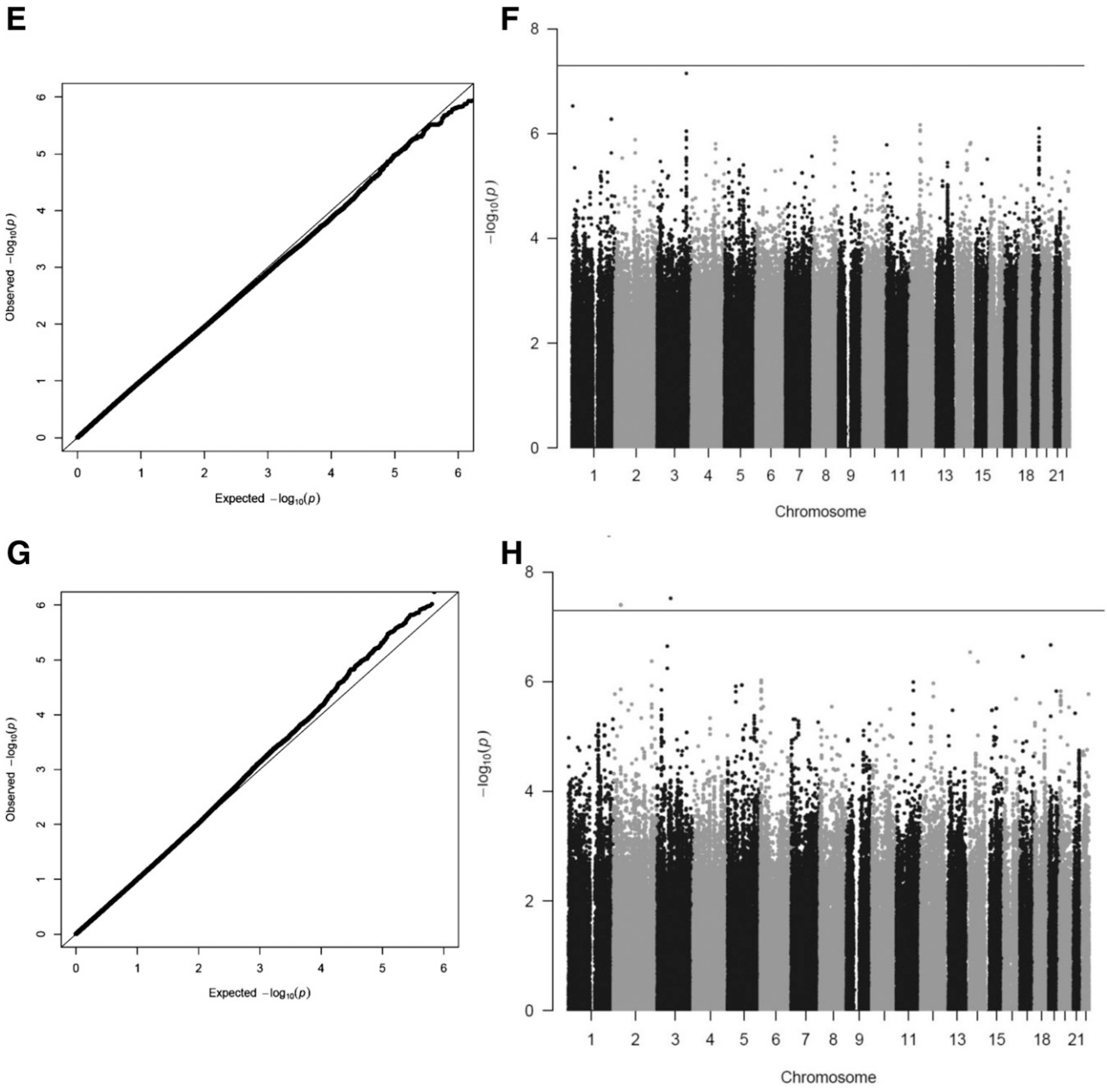


**Figure 1**—*Continued*.

significance in African Americans. We could not reproduce this finding in the replication cohorts. This variant is located near the *GOLIM4* gene, which helps process proteins and mediates protein transport. The SNP rs115523882 specifically changes a motif that is a binding site for Nlx3, a transcription factor in blood, suggesting it plays a regulatory role. This variant is mainly present in people of African ancestry (minor allele frequency [MAF] = 0.0393) and not common in other ethnic groups, suggesting we may have had insufficient power to replicate it.

Of note, there was one SNP, rs184340784, suggestively associated with DR ($P = 3.52 \times 10^{-8}$) in the extremes of DR analysis without covariates in African Americans that was not present in our replication cohorts (due to low MAF) and thus could not be replicated. Neither rs115523882 nor rs184340784 was analyzed for eQTL activity in GTEx due to their low MAF (MAF $< 0.01$ in GTEx tissues).

Table 6 and Supplementary Table 29 show the discovery variants with $P < 1 \times 10^{-5}$ that achieved a nominal $P < 0.05$ in the complete replication sample or in one of the replication ethnicities, respectively, and had the same direction as the discovery samples. None of these variants achieved genome-wide significance after discovery and replication meta-analysis, as defined above.

**Table 5—Replication results for variants with $P < 5 \times 10^{-8}$ (traditional, nominal threshold for genome-wide significance) in the discovery analysis**

| Discovery population/LT modeling | RSID | Nearest gene | REF | Disc NEFF | Disc RAF | Disc $P$ | Disc OR | All rep NEFF | All rep RAF | All rep OR | All rep $P$ | Disc + rep OR (95% CI) | Disc + rep $P$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variants identified in the PDR discovery analysis | | | | | | | | | | | | | |
| AA/no | rs115523882 | *GOLIM4* | A | 1,452 | 0.9721 | $9.42 \times 10^{-9}$ | 3.10 | 571 | 0.9975 | 0.20 | 0.13 | 2.89 (1.97, 4.23) | $8.51 \times 10^{-8}$ |
| AA/yes | rs115523882 | *GOLIM4* | A | 1,452 | 0.9721 | $5.37 \times 10^{-9}$ | 3.10 | 571 | 0.9975 | 0.20 | 0.18 | 2.89 (1.99, 4.20) | $4.25 \times 10^{-8}$ |
| European/no | rs139205645 | *NDUFB3* | T | 907 | 0.9907 | $3.93 \times 10^{-8}$ | 0.13 | 3,431 | 0.9900 | 0.74 | 0.77 | 0.48 (0.29, 0.79) | 0.004 |
| European/yes | rs17791,488 | *NOS2/LYRM9* | T | 907 | 0.9705 | $7.26 \times 10^{-9}$ | 3.70 | 5,883 | 0.9772 | 0.82 | 0.33 | 1.08 (0.98, 1.19) | 0.12 |
| Variants identified in the extremes of DR analysis | | | | | | | | | | | | | |
| AA/no | rs184340784 | *AJAP1* | C | 603 | 0.0063 | $3.52 \times 10^{-8}$ | NA | * | * | * | * | — | — |
| European/yes | rs142293996 | *NVL* | C | 523 | 0.9895 | $2.10 \times 10^{-9}$ | 2.38 | 1,229 | 0.9910 | 3.23 | 0.16 | 2.91 (1.85, 4.57) | $4.10 \times 10^{-6}$ |
| European/yes | rs17706,958 | *PDZRN3* | T | 797 | 0.7615 | $3.04 \times 10^{-8}$ | 1.58 | 4,194 | 0.9828 | 1.28 | 0.02 | 1.39 (1.24, 1.56) | $7.41 \times 10^{-8}$ |
| European/yes | rs80117617 | *SLC8A1* | T | 797 | 0.9598 | $4.04 \times 10^{-8}$ | 3.78 | 3,345 | 0.9726 | 1.29 | 0.24 | 1.71 (1.30, 2.25) | $1.35 \times 10^{-4}$ |

AA, African American; All rep, all replication cohorts; CHR, chromosome; Disc, discovery; LT, liability threshold; NA, not available; NEFF, effective sample size; RAF, reference allele frequency in sample; REF, reference allele; Rep, replication; RSID, rs identifier. *None of the replication cohorts were able to provide data for this SNP.

### DAPPLE Results: Protein–Protein Interactions

One protein network from the African American PDR analysis was significant ($P = 0.0009$) for average binding degree within the network (Fig. 2). The aforementioned top-ranked SNP (rs115523882) could not be included in the DAPPLE analysis because its nearby gene (*GOLIM4*) is not in the protein database. The significant protein network includes genes with primary roles in inflammation including *IFNG*, *IL22RA1*, *CFH*, and *SELL*. *IFNG* encodes interferon-γ, which is highly expressed in ocular tissues from patients with PDR (43). *IL22RA1* encodes the IL-22 receptor, and *CFH* encodes complement factor H; both proteins are suspected to play a role in PDR (44,45). *SELL* encodes L-selectin, which is expressed at higher levels in lymphocytes from patients with DR and associated with increased endothelial adhesion (46). We did not identify any statistically significant protein networks for any of the other case-control definitions in African Americans or Europeans.

### MAGENTA Confirmation of DAPPLE Results

We examined the 41 genes in the significant network identified by the DAPPLE analysis via GSEA using MAGENTA. The genes showed a significant (16.5-fold) enrichment of low association $P$ values in the African American PDR analysis ($P < 1 \times 10^{-6}$) (Supplementary Fig. 7 and Supplementary Table 30) and to a lesser extent in African American extremes of DR analysis ($P = 2 \times 10^{-4}$) (Supplementary Table 30), suggesting new DR associations of modest effects in African Americans (Supplementary Table 31). No significant gene set enrichment was found for the PDR and extremes of DR analyses in Europeans.

### Loci Associated With Type 2 Diabetes and Glycemic Traits

The results of the correlation analysis between type 2 diabetes/glycemic trait-associated SNPs and DR case status are shown in Supplementary Table 32. The Z score for type 2 diabetes was +2.256 ($P = 0.024$). The correlation coefficient $R$ was positive, indicating that a greater burden of SNPs that increase type 2 diabetes risk is correlated with having DR. However, this Z score was not significant after correcting for the six hypotheses (six traits) tested.

### Previously Associated SNPs From Prior Studies

We extracted results from our discovery meta-analysis for the variants with the lowest association $P$ values from previously published DR GWAS or large candidate gene studies (Supplementary Table 33). There were three variants that were nominally significant ($P < 0.05$) in our sample and had the same direction of effect as in the previously published studies. Two of the variants, rs9896052 and rs6128, were from previous studies for which samples overlapped with some samples in our discovery meta-analysis and therefore do not represent

**Table 6—Replication results for variants with nominal significance ($P < 0.05$) in the combined (Hispanic, African American, and European cohorts) replication meta-analyses**

| Discovery population/LT modeling | RSID | Nearest gene | REF* | Disc EAF | Disc OR | Disc *P* | All rep OR | All rep *P* | Disc + rep OR | Disc + rep *P* |
|---|---|---|---|---|---|---|---|---|---|---|
| Variants identified in the any DR discovery analysis | | | | | | | | | | |
| European (Sens)/no | rs1394919 | *PPEF2/NAAA* | C | 0.72 | 0.73 | $8.51 \times 10^{-6}$ | 0.91 | 0.003 | 0.88 | $6.35 \times 10^{-6}$ |
| AA (Sens)/no | rs75360147 | *SLC28A3* | T | 0.93 | 2.08 | $7.07 \times 10^{-6}$ | 2.65 | 0.009 | 2.17 | $2.29 \times 10^{-7}$ |
| European/no | rs1508244 | *HTR1E* | A | 0.98 | 0.33 | $3.74 \times 10^{-6}$ | 0.92 | 0.01 | 0.90 | 0.002 |
| ME/no | rs10432638 | *UBXN2A* | C | 0.73 | 0.78 | $2.60 \times 10^{-6}$ | 0.93 | 0.01 | 0.89 | $7.74 \times 10^{-6}$ |
| EU/no | rs150775408 | *BC031225* | C | 0.95 | 1.97 | $7.24 \times 10^{-6}$ | 1.27 | 0.04 | 1.46 | $2.54 \times 10^{-5}$ |
| AA/yes | rs143894698 | *GCM1* | G | 0.98 | 3.14 | $4.62 \times 10^{-6}$ | 1.45 | 0.004 | 1.58 | $2.53 \times 10^{-5}$ |
| European/yes | rs13006587 | *ATAD2B* | G | 0.58 | 0.79 | $7.52 \times 10^{-6}$ | 0.93 | 0.006 | 0.92 | $4.74 \times 10^{-5}$ |
| European/yes | rs73642012 | *PTPRD* | C | 0.91 | 0.67 | $9.58 \times 10^{-6}$ | 0.90 | 0.02 | 0.87 | $8.67 \times 10^{-5}$ |
| Variants identified in the PDR discovery analysis | | | | | | | | | | |
| Europeans/no | rs139921826 | *PRSS35* | G | 0.98 | 0.33 | $7.92 \times 10^{-6}$ | 0.66 | 0.03 | 0.62 | 0.0008 |
| AA/yes | rs1414474 | *C1orf94* | C | 0.14 | 1.62 | $1.46 \times 10^{-7}$ | 1.12 | 0.01 | 1.19 | $1.90 \times 10^{-5}$ |
| AA/yes | rs9998354 | *BTF3P13* | T | 0.44 | 0.73 | $8.74 \times 10^{-6}$ | 0.92 | 0.04 | 0.87 | 0.0001 |
| European/yes | rs142293996 | *NVL* | C | 0.99 | 1.83 | $1.14 \times 10^{-6}$ | 2.40 | 0.04 | 2.29 | 0.0001 |
| Variants identified in the NPDR discovery analysis | | | | | | | | | | |
| European/no | rs1508244 | *RN7SL643P* | A | 0.98 | 0.32 | $8.13 \times 10^{-6}$ | 0.89 | 0.005 | 0.87 | 0.0005 |
| European/no | rs7944308 | *KCNA4* | G | 0.42 | 0.71 | $7.76 \times 10^{-7}$ | 0.94 | 0.02 | 0.90 | $5.80 \times 10^{-5}$ |
| Variants identified in the extremes of DR discovery analysis | | | | | | | | | | |
| AA/no | rs74161190 | *TCERG1L* | A | 0.94 | 0.32 | $4.57 \times 10^{-6}$ | 0.40 | 0.03 | 0.32 | $7.16 \times 10^{-7}$ |
| European/yes | rs17706958 | *PDZRN3* | T | 0.76 | 1.58 | $3.04 \times 10^{-8}$ | 1.28 | 0.02 | 1.39 | $7.41 \times 10^{-8}$ |
| European/yes | rs10932347 | *CPS1* | A | 0.04 | 0.33 | $4.22 \times 10^{-7}$ | 0.64 | 0.02 | 0.55 | $1.30 \times 10^{-5}$ |
| AA/yes | rs2690028 | *KAZN* | C | 0.32 | 0.62 | $4.52 \times 10^{-6}$ | 0.80 | 0.03 | 0.74 | $1.72 \times 10^{-5}$ |
| European/yes | rs116972715 | *DSC3* | C | 0.99 | 2.60 | $2.48 \times 10^{-6}$ | 3.62 | 0.03 | 3.29 | $1.59 \times 10^{-5}$ |
| European/yes | rs75167957 | *CTNNA2* | C | 0.99 | 3.26 | $3.36 \times 10^{-6}$ | 9.77 | 0.04 | 6.34 | $5.83 \times 10^{-6}$ |
| AA/yes | rs6577631 | *LOC339862* | G | 0.86 | 0.53 | $3.45 \times 10^{-6}$ | 0.89 | 0.04 | 0.84 | 0.0006 |

AA, African American; All rep, all replication cohorts; Disc, discovery; ME, multiethnic; REF, reference allele; Rep, replication; Sens, sensitivity analysis. *For insertion-deletions, the reference allele is shown first followed by the alternate allele.

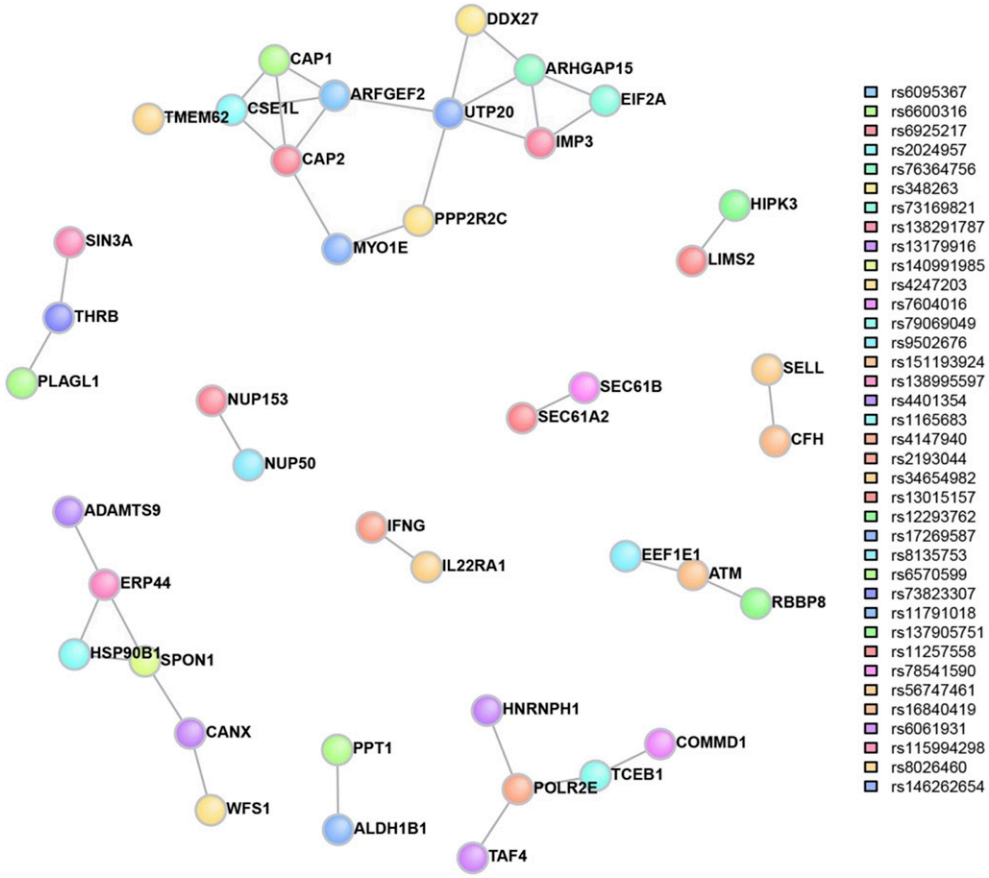


**Figure 2**—Protein network from the African American PDR discovery analysis that was significant in the DAPPLE analysis. This significant protein network includes genes with primary roles in inflammation (*IFNG*, *IL22RA1*, *CFH*, and *SELL*), protein function/endoplasmic reticulum function (*ADAMT30*, *ERP44*, *HSP90B1*, *SPON1*, *CNAX*, and *WFS1*), catabolic processing/metabolism (*PPT1* and *ALDH1B1*), gene expression/transcription factor activity (*HNRNPH1*, *TAF4*, *POLR2E*, *TCEB1*, *COMMD1*, *PLAGL1*, *THRB*, and *SIN3A*), macromolecule transport (*NUP153* and *NUP50*), protein localization (*SEC61B* and *SEC61A2*), and DNA repair/cell cycle (*RBBP8*, *ATM*, and *EEF1E1*).

independent replication (10,20). Variant rs1399634, originally found in Chinese patients ($P = 2 \times 10^{-6}$), was nominally significant in our European discovery cohort ($P = 0.0124$). Meta-analysis of the original study and our cohorts was performed using the same method as our discovery and replication meta-analyses and was short of genome-wide significance (OR 1.47; $P = 9.63 \times 10^{-8}$).

## DISCUSSION

To our knowledge, this study represents the largest GWAS performed for DR. The discovery analysis included 3,246 Europeans and 2,611 African Americans. The replication analysis included 18,545 Europeans, 16,453 Asians, and 2,710 Hispanics. Despite the relatively large sample size, we did not identify any individual variants that were associated at a genome-wide significant level after meta-analysis with multiethnic replication cohorts. However, among the most significant results in the African American PDR analysis, we did identify a statistically significant enrichment for a network of genes using DAPPLE, which was corroborated by GSEA using MAGENTA.

In the discovery meta-analyses, several variants from the PDR and extremes of DR analyses achieved nominal genome-wide significance of $P < 5 \times 10^{-8}$, but the only variant to achieve genome-wide significance after conservative multiple testing correction was rs142293996 in the European analysis for extremes of DR ($P = 2.1 \times 10^{-9}$). It is notable that the variants with the most significant findings came from the two case-control definitions that have PDR as their case definition. This is consistent with the fact that PDR has a higher heritability than overall DR (4). Although the most strongly associated variants in the discovery analyses (rs142293996 in *NVL* in Europeans and rs115523882 in *GOLIM4* in African Americans) did not reach genome-wide significance with replication, it is still possible that they do play a role in DR pathogenesis. *NVL* is highly expressed in the retina, and the implicated variant is in LD with an eQTL acting on *CAPN2* with functional implications in neural tissue. The eQTL variant falls in a binding site of a transcription factor (47). The *GOLIM4* variant also has a known regulatory role.

We could not replicate the association with rs142293996 when we used the Fisher exact test, although the Fisher

exact test did not allow for covariate incorporation. There is potential for inflated false-positive rate when standard association methods are applied to rare (e.g., MAF <1%) variants in imbalanced (e.g., case fraction <10%) case-control cohorts at modest sample sizes (48). However, most cohorts in this study did not have case fraction <10%. Larger sample sizes will help determine the confidence in these top associations.

There was one variant suggestively associated in the extremes of DR discovery analysis in African Americans, rs184340784, which was not present in any replication data sets. The T allele of this variant has a frequency of 0.0023 in African populations and 0 in European, East Asian, South Asian, and Hispanic populations in the 1000 Genomes phase 3 panel. In the discovery analysis, the $P = 3.52 \times 10^{-8}$ was shy of the genome-wide significance threshold of $3.75 \times 10^{-9}$ for variants discovered from the African ancestry analyses. This variant is within an intronic region upstream of adherens junctions–associated protein 1 (*AJAP1*), which has its highest expression in brain frontal cortex but is also expressed in the retina (https://www.proteinatlas.org/ENSG00000196581-AJAP1/tissue).

In the DAPPLE analysis, we did find that the top signals for the PDR analyses in African Americans analysis were enriched for a biologic network. The advantage of DAPPLE is that it can identify a protein pathway that may not be evident solely from the primary individual variant GWAS. The presence of an underlying network among the top loci suggests there are likely true associations within top findings that have yet to reach genome-wide significance due to limited power. Multiple pathways including inflammatory pathways are implicated by this network. To confirm biological significance, these results will need to be followed up with functional in vitro studies.

The DAPPLE results were corroborated by the MAGENTA GSEA in the African American PDR and extremes of DR analyses. This network of genes, however, was not enriched for in Europeans. This could either be due to technical differences (e.g., the number of African American cases is approximately threefold larger than the number of European cases) or due to biological reasons. For example, we found that the allele frequencies of the most significant variant per gene for 40% of these protein-interacting genes are rare in Europeans (MAF <0.2%), whereas they are common in African Americans (MAF >1%), according to the Genome Aggregation Database (see Web Resources in the Supplemental Data).

In the analysis between type 2 diabetes/glycemic trait SNPs and DR case status, only type 2 diabetes variants were significantly associated with DR prior to, but not after, multiple testing correction. One previous study examined aggregate effects of 76 type 2 diabetes–associated variants in Asian patients (49). Participants in the top tertile of type 2 diabetes risk score were 2.56-fold more likely to have DR compared with lowest tertile participants. Our study's result showed the same direction of effect as in the prior study, with type 2 diabetes risk-raising alleles increasing DR risk. The prior study did not examine glycemic traits. Our inability to detect a correlation for glycemic traits may be due to the small amount of glycemic variance captured by these variants. In European patients, $HbA_{1c}$ SNPs explain ~5% of $HbA_{1c}$ variance (50).

We were unable to replicate findings from previous studies (6–8,12,13,17,18). We did have the same direction of effect in our European discovery sample for rs1399634 (*LRP2*), which was initially reported in an Asian population. However, the meta-analysis was shy of genome-wide significance. The overall lack of replication of previous reports' findings is not surprising, given the heterogeneity in phenotyping, case-control definitions, ethnicities, and analytic approaches, although we did try to match our case-control definitions to the original studies' definitions.

There are many potential reasons why we were unable to identify replicable, significant associations from our discovery GWAS. First, the genetic risk in DR development may be quite small in proportion to the nongenetic risk factors. Therefore, even though we assembled the largest sample, it may not be sufficient to detect very modest effects. There was heterogeneity between the discovery and replication cohorts that could contribute to inability to replicate. The discovery cohort included individuals with type 2 diabetes, whereas the replication cohorts included individuals with either type 1 or type 2 diabetes. It is not known definitively whether genetic variants for DR differ between type 1 and type 2 diabetes. Clinically, DR phenotypes are similar in patients with type 1 and type 2 diabetes, so we hypothesize that at least some of the genetic risk is shared. However, we cannot be certain of this, and heterogeneity of diabetes type might have contributed to lack of replication. The discovery cohort included individuals who were of either European or African American descent, whereas the replication cohorts included individuals of European, Hispanic, or Asian descent. This heterogeneity could also have led to lack of replication. Europeans were represented in both the discovery and replication phases, but even our European discovery analysis has limited power. Power calculations show that our discovery GWAS for the any DR analysis in Europeans had 100% power to detect a variant with an MAF of 0.40 with a heterozygous genotypic relative risk of 1.5 with a $P$ value $< 5 \times 10^{-8}$, whereas the power decreases to 5% for the same variant with genotypic relative risk of 1.2.

We attempted to harmonize the phenotypes as much as possible, but there were some limits to complete harmonization, particularly for cohorts with limited-field or no photography. Misclassification of participants because of limited DR ascertainment could have biased the results to the null. Although we did use LTSCORE modeling to account for DoD, we may have had some misclassification bias because we did not have a minimum DoD for control

subjects (i.e., some control subjects could have developed DR with longer DoD), which would also bias our result toward the null. We only had one $HbA_{1c}$ measure. Repeated $HbA_{1c}$ measures would reflect long-term glycemia more accurately.

In summary, we have executed the largest GWAS of DR to date. There were no genome-wide significant findings, but analysis of protein–protein interaction networks point to possible candidate pathways for PDR in African Americans. Future studies examining DR genetics would benefit from a greater international collaboration encompassing larger samples that would allow strict case-control definitions that define a minimal DoD without sacrificing power. Furthermore, these studies should focus case definitions on the advanced forms of DR—PDR and diabetic macular edema—and incorporate more refined phenotyping, particularly optical coherence tomography for diabetic macular edema. Finally, whole-genome sequencing might reveal a role for very rare variants, particularly for the DR phenotypic extremes.

**Acknowledgments.** The authors thank the staff at the Icelandic Heart Association and the Age, Gene/Environment Susceptibility - Reykjavik Study (AGES) participants who volunteered time and allowed the authors to contribute data to this international project; the staff and participants of the Atherosclerosis Risk in Communities (ARIC) study for important contributions; and all of the patients recruited in the Genetics of Diabetes and Audit Research Tayside Study (GoDARTS) and other European and African American cohorts. The authors especially thank the Health Informatics Centre in the School of Medicine, University of Dundee, for help with data access; the staff and participants of the Jackson Heart Study (JHS); and the other investigators, staff, and participants of the Multiethnic Study of Atherosclerosis (MESA) study for valuable contributions (a full list of participating MESA investigators and institutions can be found at https://www.mesa-nhlbi.org/). A full list of principal Cardiovascular Health Study (CHS) investigators and institutions can be found at CHS-NHLBI.org.

**Funding.** This study has received support from the following organizations for this research: Research to Prevent Blindness, Inc., National Eye Institute (EY-16335, EY-22302, EY-11753, and R01-EY-023644 and core grant EY-001792), National Institutes of Health (NIH) (R01-HG-006399), Massachusetts Lions Eye Research Fund, Alcon Research Institute, American Diabetes Association (1-11-CT-51), and Harvard Catalyst. AGES was supported by the NIH through the Intramural Research Program of the National Institute on Aging (ZIAAG007380) and the National Eye Institute (ZIAEY00401), NIH contract number N01-AG-1-2100, Hjartavernd (the Icelandic Heart Association), the Althingi (Icelandic Parliament), and the University of Iceland Research Fund.

The funders had no role in collection, management, analysis, or interpretation of data and were not involved in the preparation, writing, or approval of the article or the decision to submit the article for publication.

ARIC is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute (NHLBI) contracts HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C and R01-HL-087641, R01-HL-59367, and R01-HL-086694; National Human Genome Research Institute contract U01-HG-004402; and NIH contract HHSN268200625226C. ARIC has been funded in whole or in part with federal funds from the National Heart, Lung, and Blood Institute, NIH, Department of Health and Human Services (contract numbers HHSN2682017000001I, HHSN2682017000002I, HHSN2682017000003I, HHSN2682017000004I, and HHSN2682017000005I). Infrastructure was partly supported by the NIH and NIH Roadmap for Medical Research (grant UL1-RR-025005). Funding support for "Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium" was provided by the NIH through the American Recovery and Reinvestment Act of 2009 (5RC2HL102419). The Australian Genetics of Diabetic Retinopathy Study was supported by the Australian National Health and Medical Research Council (NHMRC) (595918), Canberra, Australia, and the Ophthalmic Research Institute of Australia. J.E.C. is supported by a Practitioner Fellowship from the NHMRC. K.P.B. is supported by a Senior Research Fellowship from the NHMRC. The Blue Mountains Eye Study (BMES) was supported by the Australian NHMRC (project grant identification numbers 974159, 211069, and 302068) and Centre for Clinical Research Excellence in Translational Clinical Research in Eye Diseases (grant 529923). The BMES GWAS and genotyping costs were supported by the Australian NHMRC (project grant identification numbers 512423, 475604, and 529912) and the Wellcome Trust (U.K.) as part of the Wellcome Trust Case Control Consortium 2 (WTCCC2) (085475/B/08/Z and 085475/08/Z). The Cardiovascular Health Study was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01-HC-55222, N01-HC-85079, N01-HC-85080, N01-HC-85081, N01-HC-85082, N01-HC-85083, N01-HC-85086, and N01-HC-75150 and grants U01-HL-080295 and U01-HL-130114, with additional contribution from the National Institute of Neurological Disorders and Stroke. Additional support was provided by National Institute on Aging grant R01-AG-023629.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Subjects included in the present analysis consented to the use of their genetic information.

This work was supported by Tenovus Scotland (2015 T15/40 to W.M.). GoDARTS was supported by Chief Scientist Office Scotland and Diabetes UK. The genotyping costs were granted by the Wellcome Trust for WTCCC2 samples and by the Innovative Medicines Initiative for SUMMIT (SUrrogate markers for Micro- and Macrovascular hard endpoints for Innovative diabetes Tools) samples. The SUMMIT consortium was supported by the European Union's Seventh Framework Programme (FP7/2007-2013) for the Innovative Medicine Initiative under grant agreement IMI/115006 (to the SUMMIT consortium). FinnDiane was supported by grants from the Folkhälsan Research Foundation, the Wilhelm and Else Stockmann Foundation, the Liv och Hälsa Foundation, Helsinki University Central Hospital Research Funds (EVO), the Novo Nordisk Foundation (NNF14SA0003), and the Academy of Finland (134379, 275614, and 299200). The JHS is supported by the NHLBI and the National Institute on Minority Health and Health Disparities and is conducted in collaboration with Jackson State University (contracts HHSN268201300049C and HHSN268201300050C), Tougaloo College (HHSN268201300048C), and the University of Mississippi Medical Center (HHSN268201300046C and HHSN268201300047C).

The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the NHLBI, the NIH, or the U.S. Department of Health and Human Services.

MESA and the MESA SHARe (SNP Health Association Resource) project are conducted and supported by the NHLBI in collaboration with MESA investigators. Support for MESA is provided by NHLBI contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, UL1-TR-001881, and DK-063491. Additional funding is provided by the Intramural Research Program of the National Eye Institute (ZIAEY000403). Funding for SHARe genotyping was provided by NHLBI contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, CA) and the Broad Institute of MIT and Harvard (Boston, MA) using the Affymetrix Genome-Wide Human SNP Array 6.0. The Singapore Epidemiology of Eye Diseases study was supported by the National Medical Research Council (NMRC), Singapore (grants 0796/2003, 1176/2008, 1149/2008, STaR/0003/2008, 1249/2010, CG/SERI/2010, CIRG/1371/2013, and CIRG/1417/2015), and Biomedical Research Council, Singapore (08/1/35/19/550 and 09/1/35/19/616). C.-Y.C. is supported by an award from the NMRC (CSA-SI/0012/2017).

NMRC had no role in the design or conduct of this research.

The Starr County Health Studies were supported, in part, by the State of Texas and NIH grants EY-012386, DK-047487, and DK-073541. The Korean Study of Diabetic Retinopathy was supported by National Research Foundation of Korea grants funded by the Korea government (NRF-2017R1A2B2011436 and NRF-2012R1A1A2008943). For the Wake Forest School of Medicine Study (WFU), genotyping services were provided by the Center for Inherited Disease Research. The Center for Inherited Disease Research is fully funded through a federal contract from the NIH to Johns Hopkins University, contract HHSC268200782096C. This work was supported by NIH grants R01-DK-087914, R01-DK-066358, R01-DK-053591, DK-070941, and DK-084149, Wake Forest School of Medicine grant M01-RR-07122, and Venture Fund. Work for this manuscript was supported in part by the Genetics of Latinos Diabetic Retinopathy (GOLDR) Study grant EY14684. This study was supported by the National Eye Institute of the NIH (EY-014684 to Y.-D.I.C. and J.I.R.) and ARRA (American Recovery and Reinvestment Act of 2009) Supplement (EY014684-03S1 and EY014684-04S1), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) DK-063491 to the Southern California Diabetes Endocrinology Research Center, the Eye Birth Defects Foundation Inc., the National Science Council, Taiwan (NSC 98-2314-B-075A-002-MY3 to W.H.-H.S.), and the Taichung Veterans General Hospital, Taichung, Taiwan (TCVGH-1003001C to W.H.-H.S.). The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1-TR-001881 and NIDDK Diabetes Research Center grant DK-063491 to the Southern California Diabetes Research Center. The Family Study of Nephropathy and Diabetes (FIND) study was supported by grants U01-DK-57292, U01-DK-57329, U01-DK-057300, U01-DK-057298, U01-DK-057249, U01-DK-57295, U01-DK-070657, U01-DK-057303, and U01-DK-57304 and, in part, by the Intramural Research Program of the NIDDK. Support was also received from NHLBI grants U01-HL-065520, U01-HL-041654, and U01-HL-041652. This project has been funded in whole or in part with federal funds from the National Cancer Institute, NIH, under contract N01-CO-12400 and the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. This work was also supported by the National Center for Research Resources for the General Clinical Research Center grants: Case Western Reserve University, M01-RR-000080; Wake Forest University, M01-RR-07122; Los Angeles Medical Center, Harbor-University of California, M01-RR-00425; College of Medicine, University of California, Irvine, M01-RR-00827–29; Health Sciences Center, University of New Mexico, M01-RR-00997; and the Frederic C. Bartter Award, M01-RR-01346. Computing resources were provided, in part, by the Wake Forest School of Medicine Center for Public Health Genomics.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

M.I.M. is a Wellcome Senior Investigator supported by Wellcome awards 090532, 106130, 098381, and 203141 and a National Institute for Health Research Senior Investigator.

The views expressed in this article are those of the authors and not necessarily those of the National Health Service, National Institute for Health Research, or the Department of Health.

**Duality of Interest.** J.Z.K. is employed by Sun Pharmaceutical Industries, Inc. P.-H.G. has received investigator-initiated research grants from Eli Lilly and Company and Roche, is an advisory board member for AbbVie, AstraZeneca, Boehringer Ingelheim, Cebix, Eli Lilly and Company, Janssen, Medscape, Merck Sharp & Dohme, Novartis, Novo Nordisk, and Sanofi, and has received lecture fees from AstraZeneca, Boehringer Ingelheim, Eli Lilly, ELO Water, Genzyme, Merck Sharp & Dohme, Medscape, Novo Nordisk, and Sanofi. B.L.Y. is a full-time employee of Genentech, Inc., and holds stock and stock options in Roche. No other potential conflicts of interest relevant to this article were reported.

Sun Pharmaceutical Industries, Inc., is not in any way involved in this study.

**Author Contributions.** S.P., E.J.R., A.V.Se., S.D., L.K.S., A.L.P., and L.S. contributed to the writing of the manuscript. R.P.I., R.A.J., M.C., X.L., C.-Y.C., M.C.Y.N., A.V.Sm., G.S.T., Y.-D.I.C., J.Z.K., L.M.D., W.M., S.M.H., M.I., D.N., J.K., Y.H., Y.J., J.A., A.L., K.S., K.H.P., X.G., E.I., K.D.T., S.G.A., J.R.S., B.I.F., I.-T.L., W.H.-H.S., M.K., A.T., S.H., M.M., D.-A.T., R.M.-C., R.V., M.I.M., L.G., E.A., V.L., E.A., A.M., A.S.F.D., H.M.C., I.T., N.S., P.-H.G., S.M., C.L.H., A.P., C.J.C., H.H., P.M., J.E.C., E.Y.C., A.D.P., M.A.G., C.P., D.W.B., B.L.Y., D.S., M.F.C., J.J.W., K.P.B., T.Y.W., B.E.K.K., R.K., J.I.R., and S.K.I. reviewed and edited the manuscript. S.P., R.P.I., R.A.J., C.-Y.C., M.C.Y.N., A.V.Sm., G.S.T., Y.-D.I.C., J.Z.K., W.M., S.M.H., M.I., J.K., J.A., A.L., K.S., K.H.P., X.G., B.I.F., S.H., M.M., D.-A.T., R.M.-C., I.T., N.S., P.-H.G., S.M., C.L.H., A.P., C.J.C., H.H., P.M., J.E.C., E.Y.C., A.D.P., M.A.G., C.P., D.W.B., B.L.Y., D.S., M.F.C., J.J.W., K.P.B., T.Y.W., B.E.K.K., R.K., J.I.R., S.K.I., A.L.P., and L.S. collected and researched data. S.P., R.P.I., R.A.J., M.C., X.L., C.-Y.C., M.C.Y.N., A.V.Sm., E.J.R., A.V.Se., S.D., G.S.T., Y.-D.I.C, J.Z.K., L.M.D., L.K.S., W.M., S.M.H., M.I., D.N., J.K., Y.H., Y.J., J.A., A.L., K.S., S.H., M.M., I.T., N.S., B.L.Y., K.P.B., A.L.P., and L.S. performed the analysis. L.S. is the guarantor of this work and, as such, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Prior Presentation.** Parts of this study were presented in abstract form at the 42nd Annual Meeting of the Macula Society, Bonita Springs, FL, 13–16 February 2019.

## References

1. Yau JW, Rogers SL, Kawasaki R, et al.; Meta-Analysis for Eye Disease (META-EYE) Study Group. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care 2012;35:556–564

2. Klein R, Klein BE, Moss SE, Davis MD, DeMets DL. The Wisconsin epidemiologic study of diabetic retinopathy. III. Prevalence and risk of diabetic retinopathy when age at diagnosis is 30 or more years. Arch Ophthalmol 1984;102:527–532

3. Looker HC, Nelson RG, Chew E, et al. Genome-wide linkage analyses to identify Loci for diabetic retinopathy. Diabetes 2007;56:1160–1166

4. Arar NH, Freedman BI, Adler SG, et al.; Family Investigation of Nephropathy and Diabetes Research Group. Heritability of the severity of diabetic retinopathy: the FIND-Eye study. Invest Ophthalmol Vis Sci 2008;49:3839–3845

5. Sheu WH, Kuo JZ, Lee IT, et al. Genome-wide association study in a Chinese population with diabetic retinopathy. Hum Mol Genet 2013;22:3165–3173

6. Awata T, Yamashita H, Kurihara S, et al. A genome-wide association study for diabetic retinopathy in a Japanese population: potential association with a long intergenic non-coding RNA. PLoS One 2014;9:e111715

7. Fu YP, Hallman DM, Gonzalez VH, et al. Identification of diabetic retinopathy genes through a genome-wide association study among Mexican-Americans from Starr County, Texas. J Ophthalmol 2010;2010:861291

8. Grassi MA, Tikhomirov A, Ramalingam S, Below JE, Cox NJ, Nicolae DL. Genome-wide meta-analysis for severe diabetic retinopathy. Hum Mol Genet 2011;20:2472–2481

9. Huang YC, Lin JM, Lin HJ, et al. Genome-wide association study of diabetic retinopathy in a Taiwanese population. Ophthalmology 2011;118:642–648

10. Burdon KP, Fogarty RD, Shen W, et al. Genome-wide association study for sight-threatening diabetic retinopathy reveals association with genetic variation near the GRB2 gene. Diabetologia 2015;58:2288–2297

11. Shtir C, Aldahmesh MA, Al-Dahmash S, et al. Exome-based case-control association study using extreme phenotype design reveals novel candidates with protective effect in diabetic retinopathy. Hum Genet 2016;135:193–200

12. Hosseini SM, Boright AP, Sun L, et al.; DCCT/EDIC Research Group. The association of previously reported polymorphisms for microvascular complications in a meta-analysis of diabetic retinopathy. Hum Genet 2015;134:247–257

13. Grassi MA, Tikhomirov A, Ramalingam S, et al. Replication analysis for severe diabetic retinopathy. Invest Ophthalmol Vis Sci 2012;53:2377–2381

14. Peng D, Wang J, Zhang R, et al. Common variants in or near ZNRF1, COLEC12, SCYL1BP1 and API5 are associated with diabetic retinopathy in Chinese patients with type 2 diabetes. Diabetologia 2015;58:1231–1238

15. Cheung CY, Hui EY, Lee CH, et al. Impact of genetic loci identified in genome-wide association studies on diabetic retinopathy in Chinese patients with type 2 diabetes. Invest Ophthalmol Vis Sci 2016;57:5518–5524

16. Zaitlen N, Lindström S, Pasaniuc B, et al. Informed conditioning on clinical covariates increases power in case-control association studies. PLoS Genet 2012; 8:e1003032
17. Grassi MA, Mazzulla DA, Knudtson MD, et al. Patient self-report of prior laser treatment reliably indicates presence of severe diabetic retinopathy. Am J Ophthalmol 2009;147:501–504
18. Meng W, Shah KP, Pollack S, et al.; Wellcome Trust Case Control Consortium 2 (WTCCC2); Surrogate Markers for Micro- and Macro-Vascular Hard Endpoints for Innovative Diabetes Tools (SUMMIT) Study Group. A genome-wide association study suggests new evidence for an association of the NADPH Oxidase 4 (NOX4) gene with severe diabetic retinopathy in type 2 diabetes. Acta Ophthalmol 2018; 96:e811–e819.
19. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the expert committee on the diagnosis and classification of diabetes mellitus. Diabetes Care 2003;26(Suppl. 1):S5–S20
20. Sobrin L, Green T, Sim X, et al.; Family Investigation of Nephropathy and Diabetes-Eye Research Group; Wellcome Trust Case Control Consortium 2. Candidate gene association study for diabetic retinopathy in persons with type 2 diabetes: the Candidate gene Association Resource (CARe). Invest Ophthalmol Vis Sci 2011;52:7593–7602
21. Wong TY, Klein R, Islam FM, et al. Diabetic retinopathy in a multi-ethnic cohort in the United States. Am J Ophthalmol 2006;141:446–455
22. Klein R, Marino EK, Kuller LH, et al. The relation of atherosclerotic cardiovascular disease to retinopathy in people with diabetes in the Cardiovascular Health Study. Br J Ophthalmol 2002;86:84–90
23. Klein R, Sharrett AR, Klein BE, et al.; ARIC Group. The association of atherosclerosis, vascular risk factors, and retinopathy in adults with diabetes: the Atherosclerosis Risk in Communities Study. Ophthalmology 2002;109:1225–1234
24. Kaidonis G, Abhary S, Daniell M, et al. Genetic study of diabetic retinopathy: recruitment methodology and analysis of baseline characteristics. Clin Exp Ophthalmol 2014;42:486–493
25. Mitchell P, Smith W, Wang JJ, Attebo K. Prevalence of diabetic retinopathy in an older community. The Blue Mountains Eye Study. Ophthalmology 1998;105:406–411
26. Gunnlaugsdottir E, Halldorsdottir S, Klein R, et al. Retinopathy in old persons with and without diabetes mellitus: the Age, Gene/Environment Susceptibility–Reykjavik Study (AGES-R). Diabetologia 2012;55:671–680
27. Nguyen QD, Brown DM, Marcus DM, et al.; RISE and RIDE Research Group. Ranibizumab for diabetic macular edema: results from 2 phase III randomized trials: RISE and RIDE. Ophthalmology 2012;119:789–801
28. Penman A, Hoadley S, Wilson JG, Taylor HA, Chen CJ, Sobrin L. P-selectin plasma levels and genetic variant associated with diabetic retinopathy in African Americans. Am J Ophthalmol 2015;159:1152–1160.e2
29. Kuo JZ, Guo X, Klein R, et al. Systemic soluble tumor necrosis factor receptors 1 and 2 are associated with severity of diabetic retinopathy in Hispanics. Ophthalmology 2012;119:1041–1046
30. Early Treatment Diabetic Retinopathy Study Research Group. Grading diabetic retinopathy from stereoscopic color fundus photographs–an extension of the modified Airlie House classification. ETDRS report number 10. Ophthalmology 1991;98(Suppl.):786–806
31. Armitage P. Tests for linear trends in proportions and frequencies. Biometrics 1955;11:375–386
32. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–909
33. Zeggini E, Scott LJ, Saxena R, et al.; Wellcome Trust Case Control Consortium. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008;40:638–645
34. Evangelou E, Ioannidis JP. Meta-analysis methods for genome-wide association studies and beyond. Nat Rev Genet 2013;14:379–389
35. Kanai M, Tanaka T, Okada Y. Empirical estimation of genome-wide significance thresholds based on the 1000 Genomes Project data set. J Hum Genet 2016;61:861–866
36. Brown BC, Ye CJ, Price AL, Zaitlen N; Asian Genetic Epidemiology Network Type 2 Diabetes Consortium. Transethnic genetic-correlation estimates from summary statistics. Am J Hum Genet 2016;99:76–88
37. Rossin EJ, Lage K, Raychaudhuri S, et al.; International Inflammatory Bowel Disease Genetics Constortium. Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. PLoS Genet 2011;7:e1001273
38. Lundby A, Rossin EJ, Steffensen AB, et al.; QT Interval International GWAS Consortium (QT-IGC). Annotation of loci from genome-wide association studies using tissue-specific quantitative interaction proteomics. Nat Methods 2014;11: 868–874
39. Neale BM, Kou Y, Liu L, et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature 2012;485:242–245
40. Segrè AV, Wei N, Altshuler D, Florez JC, Altshuler D; DIAGRAM Consortium; MAGIC Investigators. Pathways targeted by antidiabetes drugs are enriched for multiple genes associated with type 2 diabetes risk. Diabetes 2015;64:1470–1483
41. Her J, Chung IK. The AAA-ATPase NVL2 is a telomerase component essential for holoenzyme assembly. Biochem Biophys Res Commun 2012; 417:1086–1092
42. D'Adamo CR, Dawson VJ, Ryan KA, et al. The CAPN2/CAPN8 locus on chromosome 1q is associated with variation in serum alpha-carotene concentrations. J Nutrigenet Nutrigenomics 2016;9:254–264
43. Paine SK, Basu A, Mondal LK, et al. Association of vascular endothelial growth factor, transforming growth factor beta, and interferon gamma gene polymorphisms with proliferative diabetic retinopathy in patients with type 2 diabetes. Mol Vis 2012;18:2749–2757
44. Takeuchi M, Sato T, Tanaka A, et al. Elevated levels of cytokines associated with Th2 and Th17 cells in vitreous fluid of proliferative diabetic retinopathy patients. PLoS One 2015;10:e0137358
45. Wang J, Yang MM, Li YB, Liu GD, Teng Y, Liu XM. Association of CFH and CFB gene polymorphisms with retinopathy in type 2 diabetic patients. Mediators Inflamm 2013;2013:748435
46. MacKinnon JR, Knott RM, Forrester JV. Altered L-selectin expression in lymphocytes and increased adhesion to endothelium in patients with diabetic retinopathy. Br J Ophthalmol 2004;88:1137–1141
47. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 2012;40:D930–D934
48. Zhou W, Nielsen JB, Fritsche LG, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. Nat Genet 2018;50:1335–1341
49. Chong YH, Fan Q, Tham YC, et al. Type 2 diabetes genetic variants and risk of diabetic retinopathy. Ophthalmology 2017;124:336–342
50. Wheeler E, Leong A, Liu CT, et al.; EPIC-CVD Consortium; EPIC-InterAct Consortium; Lifelines Cohort Study. Impact of common genetic determinants of hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: a transethnic genome-wide meta-analysis. PLoS Med 2017;14: e1002383